

# Cell-Free versus Cellular Massive MIMO: What Processing is Needed for Cell-Free to Win?

Emil Björnson\*, Luca Sanguinetti†

\*Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden (emil.bjornson@liu.se)

†Dipartimento di Ingegneria dell'Informazione, University of Pisa, 56122 Pisa, Italy (luca.sanguinetti@unipi.it)

**Abstract**—Communication by joint signal processing from many distributed access points, called Cell-Free Massive MIMO, is a potential beyond-5G network infrastructure. The aim of this paper is to provide the first comprehensive comparison with Cellular Massive MIMO. The uplink spectral efficiencies of four different cell-free implementations are analyzed, with spatially correlated fading and arbitrary processing. It turns out that it is possible to outperform cellular networks by a wide margin, but only using the right signal processing. A centralized implementation with optimal processing maximizes performance and, surprisingly, also reduces the fronthaul signaling.

**Index Terms**—Cell-Free Massive MIMO, cellular, uplink.

## I. INTRODUCTION

By equipping existing base stations (BSs) with Massive multiple-input multiple-output (mMIMO), the spectral efficiency (SE) can be improved by at least  $10\times$  over legacy cellular networks [1], without having to deploy new BSs. The SE gain comes from each BS having a co-located array of a hundred or more antennas, which are used for digital beamforming and, particularly, to spatially multiplex many user equipments (UEs). Signal processing methods, such as minimum mean-squared error (MMSE) combining in the uplink, can be used to cancel interference at each BS [1], without the need of any BS cooperation.

Recently, an alternative network infrastructure was proposed in [2], [3] and called *Cell-Free mMIMO*. The idea is to deploy a large number of distributed single-antenna access points (APs), which are connected to a central processing unit (CPU), also known as edge-cloud processor [4]. The latter operates the system in a Network MIMO fashion to jointly serve all the UEs by coherent joint transmission and reception [5], [6]. The paper [2] advocated the use of maximum ratio (MR) processing (a.k.a. matched filtering) at each AP, while [3], [7] showed that partially or fully centralized signal processing at the CPU can achieve higher SE. The focus in the early papers [2], [3] was on comparing Cell-Free mMIMO with a small cell network; that is, each AP serves its own exclusive set of UEs. Since small cells are a special case of Cell-Free mMIMO, they obviously provide lower performance.

In this paper, we instead compare Cell-Free mMIMO with conventional Cellular mMIMO.<sup>1</sup> These network topologies are illustrated in Fig. 1. The large differences make the comparison non-trivial and provide interesting inputs into the design of

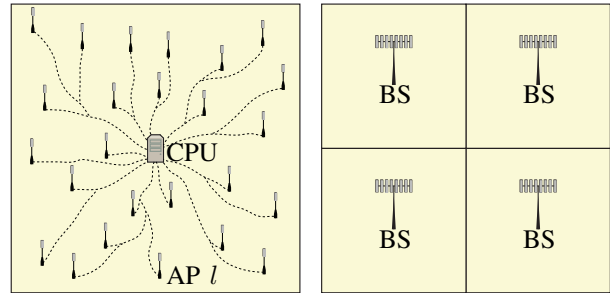


Fig. 1. Illustration of Cell-Free (left) and Cellular mMIMO (right).

beyond-5G networks. Cellular mMIMO benefits from channel hardening and spatial interference suppression, but cell-edge UEs can have bad channel conditions. Cell-Free mMIMO, on the other hand, benefits from strong macro diversity but the interference suppression capability highly depends on how it is operated. To achieve a reasonably fair comparison, we focus on the uplink with fixed UE transmit powers. The following are the major contributions of this paper:

- We describe four levels of receiver cooperation (i.e., from fully centralized to fully distributed) in Cell-Free mMIMO with multi-antenna APs. We provide new SE expressions for spatially correlated fading channels with arbitrary or optimized receive combining.
- We compare Cell-Free and Cellular mMIMO numerically and explain how to choose the right receive combining.

**Notation:** The  $n \times n$  identity matrix is  $\mathbf{I}_n$ . A block-diagonal matrix with the diagonal blocks  $\mathbf{A}_1, \dots, \mathbf{A}_n$  is denoted  $\text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_n)$ . The multi-variate circularly symmetric complex Gaussian distribution with correlation matrix  $\mathbf{R}$  is denoted  $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R})$ . The expected value is denoted  $\mathbb{E}\{\cdot\}$ .

## II. CELL-FREE NETWORK MODEL

We consider a cell-free network consisting of  $L$  geographically distributed APs, each equipped with  $N$  antennas; see Fig. 1. There are  $K$  single-antenna UEs and the channel between AP  $l$  and UE  $k$  is denoted by  $\mathbf{h}_{kl} \in \mathbb{C}^N$ . We assume that each  $\mathbf{h}_{kl}$  is constant in time-frequency blocks of  $\tau_c$  channel uses [1]. In each block, an independent realization from a correlated Rayleigh fading distribution is drawn:

$$\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_{kl}) \quad (1)$$

where  $\mathbf{R}_{kl} \in \mathbb{C}^{N \times N}$  is the spatial correlation matrix, which describes the spatial properties of the channel and  $\beta_{kl} = \text{tr}(\mathbf{R}_{kl})/N$  is the large-scale fading (including pathloss and shadowing). We assume that the channel correlation matrices  $\{\mathbf{R}_{kl} : k = 1, \dots, K\}$  are locally available at AP  $l$  [1].

E. Björnson was supported by ELLIIT and the Wallenberg AI, Autonomous Systems and Software Program (WASP). L. Sanguinetti was supported by the University of Pisa under the PRA 2018-2019 Research Project CONCEPT.

<sup>1</sup>A previous comparison is found in [8] but only for a single cell, so it is not cellular, and only MR is used, which is known to perform badly [1].

We consider the uplink, which consists of  $\tau_p$  channel uses dedicated for pilots and  $\tau_c - \tau_p$  channel uses for payload data.

#### A. Pilot Transmission and Channel Estimation

There are  $\tau_p$  mutually orthogonal  $\tau_p$ -length pilot signals that are assigned to the UEs. The case of practical interest is a large network with  $K > \tau_p$  so that more than one UE is assigned to each pilot. We denote the index of the pilot used by UE  $k$  as  $t_k \in \{1, \dots, \tau_p\}$  and call  $\mathcal{P}_k \subset \{1, \dots, K\}$  the subset of UEs that use the same pilot as UE  $k$ , including itself. When these UEs transmit their pilot, the received signal  $\mathbf{z}_{t_{kl}} \in \mathbb{C}^N$  at AP  $l$  is [1, Sec. 3]

$$\mathbf{z}_{t_{kl}} = \sum_{i \in \mathcal{P}_k} \sqrt{p_i \tau_p} \mathbf{h}_{il} + \mathbf{n}_{t_{kl}} \quad (2)$$

where  $p_i$  is the transmit power of UE  $i$ ,  $\tau_p$  is the processing gain, and  $\mathbf{n}_{t_{kl}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$  is the receiver noise with power  $\sigma^2$ . Using standard results from estimation theory [1, Sec. 3], the minimum mean-square error (MMSE) estimate of  $\mathbf{h}_{kl}$  is:

$$\hat{\mathbf{h}}_{kl} = \sqrt{p_k \tau_p} \mathbf{R}_{kl} \Psi_{t_{kl}}^{-1} \mathbf{z}_{t_{kl}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, p_k \tau_p \mathbf{R}_{kl} \Psi_{t_{kl}}^{-1} \mathbf{R}_{kl}) \quad (3)$$

where

$$\Psi_{t_{kl}} = \mathbb{E}\{\mathbf{z}_{t_{kl}} \mathbf{z}_{t_{kl}}^H\} = \sum_{i \in \mathcal{P}_k} \tau_p p_i \mathbf{R}_{il} + \mathbf{I}_N \quad (4)$$

is the correlation matrix of the received signal in (2). The estimation error  $\tilde{\mathbf{h}}_{kl} = \mathbf{h}_{kl} - \hat{\mathbf{h}}_{kl}$  is independent of  $\hat{\mathbf{h}}_{kl}$  and distributed as  $\tilde{\mathbf{h}}_{kl} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_{kl})$ , with

$$\mathbf{C}_{kl} = \mathbf{R}_{kl} - p_k \tau_p \mathbf{R}_{kl} \Psi_{t_{kl}}^{-1} \mathbf{R}_{kl}. \quad (5)$$

#### B. Data Transmission

During data transmission, UE  $k$  transmits its data signal  $s_k$  with power  $p_k$ . The received signal  $\mathbf{y}_l \in \mathbb{C}^N$  at AP  $l$  is

$$\mathbf{y}_l = \sum_{i=1}^K \mathbf{h}_{il} s_i + \mathbf{n}_l \quad (6)$$

where  $\mathbf{n}_l \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$  is the receiver noise.

### III. FOUR LEVELS OF RECEIVER COOPERATION

The APs are connected via fronthaul to a CPU; see Fig. 1. AP  $l$  receives the signal  $\mathbf{y}_l$  and can use the available channel estimates  $\{\hat{\mathbf{h}}_{kl} : k = 1, \dots, K\}$  to detect the data locally, or can fully or partially delegate this task to the CPU. The benefit of using the CPU is that it can combine the inputs from all APs. Four levels of receiver cooperation are described below and are compared with Cellular mMIMO in Section IV.

#### A. Level 4: Fully Centralized Processing

The most advanced level of Cell-Free operation is when the  $L$  APs send their received pilots  $\{\mathbf{z}_{tl} : t = 1, \dots, \tau_p, l = 1, \dots, L\}$  and data signals  $\{\mathbf{y}_l : l = 1, \dots, L\}$  to the CPU, which performs channel estimation and signal detection. Each AP needs to send  $N\tau_c$  complex scalars per coherence block. The CPU also needs to know the  $KLN^2/2$  complex scalars of the statistical matrices  $\{\mathbf{R}_{kl} : k = 1, \dots, K, l = 1, \dots, L\}$ , which are fixed for a long time; see Table I for a summary.

TABLE I  
NUMBER OF COMPLEX SCALARS TO SEND OVER FRONTHAUL TO THE CPU

	Each coherence block	Statistical parameters
Level 4	$\tau_c N L$	$KLN^2/2$
Level 3	$(\tau_c - \tau_p)KL$	$KL + (L^2 K^2 + KL)/2$
Level 2	$(\tau_c - \tau_p)KL$	—
Level 1	—	—

The received signal available at the CPU is expressed as

$$\underbrace{\begin{bmatrix} \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_L \end{bmatrix}}_{=\mathbf{y}} = \sum_{i=1}^K \underbrace{\begin{bmatrix} \mathbf{h}_{i1} \\ \vdots \\ \mathbf{h}_{iL} \end{bmatrix}}_{=\mathbf{h}_i} s_i + \underbrace{\begin{bmatrix} \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_L \end{bmatrix}}_{=\mathbf{n}} \quad (7)$$

or, in a more compact form, as

$$\mathbf{y} = \sum_{i=1}^K \mathbf{h}_i s_i + \mathbf{n}. \quad (8)$$

The collective channel is distributed as  $\mathbf{h}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_k)$  where  $\mathbf{R}_k = \text{diag}(\mathbf{R}_{k1}, \dots, \mathbf{R}_{kL}) \in \mathbb{C}^{M \times M}$  is the spatial correlation matrix, implying that (8) is mathematically equivalent to the system model of a single-cell mMIMO system with correlated fading [1, Sec. 2.3.1]. The only difference from conventional mMIMO is how the correlation matrices are generated and that multiple UEs use the same pilot, leading to pilot contamination between UEs served by the same entity.

The CPU can select an arbitrary receive combining vector  $\mathbf{v}_k \in \mathbb{C}^{NM}$  for UE  $k$  based on the collective channel estimates

$$\hat{\mathbf{h}}_k = \begin{bmatrix} \hat{\mathbf{h}}_{k1} \\ \vdots \\ \hat{\mathbf{h}}_{kL} \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, p_k \tau_p \mathbf{R}_k \Psi_{t_k}^{-1} \mathbf{R}_k) \quad (9)$$

where  $\Psi_{t_k} = \text{diag}(\Psi_{t_k1}, \dots, \Psi_{t_kL})$ . The estimation error is  $\tilde{\mathbf{h}}_k = \mathbf{h}_k - \hat{\mathbf{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{C}_k)$  with  $\mathbf{C}_k = \text{diag}(\mathbf{C}_{k1}, \dots, \mathbf{C}_{kL})$ . The capacity is then lower bounded by the following SE.

**Proposition 1.** *At Level 4, an achievable SE of UE  $k$  is*

$$\text{SE}_k^{(4)} = \left(1 - \frac{\tau_p}{\tau_c}\right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{(4)}\right) \right\} \quad (10)$$

where the instantaneous effective signal-to-interference-and-noise ratio (SINR) is

$$\text{SINR}_k^{(4)} = \frac{p_k |\mathbf{v}_k^H \hat{\mathbf{h}}_k|^2}{\sum_{i=1, i \neq k}^K p_i |\mathbf{v}_k^H \hat{\mathbf{h}}_i|^2 + \mathbf{v}_k^H \left( \sum_{i=1}^K p_i \mathbf{C}_i + \sigma^2 \mathbf{I}_{MN} \right) \mathbf{v}_k}. \quad (11)$$

*Proof:* The proof follows the same main steps as the proof of [1, Th. 4.1] and is therefore omitted. ■

The SE expression in (10) holds for any combiner  $\mathbf{v}_k$  and is a multi-antenna generalization of [7, Eq. (1)]. It can be easily computed using Monte Carlo methods, as done in Section IV. One low-complexity option is to use MR combining with  $\mathbf{v}_k = \hat{\mathbf{h}}_k$ , but the SINR-maximizing combining is of main interest since the CPU typically has high computational capability.

**Corollary 1.** The instantaneous SINR in (11) for UE  $k$  is maximized by the MMSE combining vector

$$\mathbf{v}_k = p_k \left( \sum_{i=1}^K p_i \left( \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^H + \mathbf{C}_i \right) + \sigma^2 \mathbf{I}_{MN} \right)^{-1} \hat{\mathbf{h}}_k \quad (12)$$

which leads to the maximum value

$$\text{SINR}_k^{(4)} = p_k \hat{\mathbf{h}}_k^H \left( \sum_{\substack{i=1 \\ i \neq k}}^K p_i \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^H + \sum_{i=1}^K p_i \mathbf{C}_i + \sigma^2 \mathbf{I}_{MN} \right)^{-1} \hat{\mathbf{h}}_k. \quad (13)$$

*Proof:* This result follows by utilizing that (10) is a generalized Rayleigh quotient with respect to  $\mathbf{v}_k$ . ■

The SINR-maximizing combiner in (12) is called *MMSE combining* since it also minimizes the mean-squared error  $\mathbb{E}\{|s_k - \mathbf{v}_k^H \mathbf{y}|^2 | \{\hat{\mathbf{h}}_i\}\}$ ; see [1, Sec. 4.1] for details.

### B. Level 3: Local Processing & Large-Scale Fading Decoding

Instead of sending  $\{\mathbf{y}_l : l = 1, \dots, L\}$  to the CPU, each AP can preprocess its signal by computing local estimates of the data that are then passed to the CPU for final decoding.

Let  $\mathbf{v}_{kl} \in \mathbb{C}^N$  be the combining vector that AP  $l$  selects for UE  $k$ , then its local estimate of  $s_k$  is

$$\check{s}_{kl} = \mathbf{v}_{kl}^H \mathbf{y}_l = \sum_{i=1}^K \mathbf{v}_{kl}^H \mathbf{h}_{il} s_i + \mathbf{v}_{kl}^H \mathbf{n}_l. \quad (14)$$

The vector minimizing the MSE  $\mathbb{E}\{|s_k - \mathbf{v}_{kl}^H \mathbf{y}_l|^2 | \{\hat{\mathbf{h}}_{il}\}\}$  is

$$\mathbf{v}_{kl} = p_k \left( \sum_{i=1}^K p_i \left( \hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \mathbf{C}_{il} \right) + \sigma^2 \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{kl} \quad (15)$$

which we call *Local MMSE (L-MMSE) combining* to distinguish it from Level 4. The local estimates  $\{\check{s}_{kl} : l = 1, \dots, L\}$  are sent to the CPU where they are linearly combined to obtain

$$\begin{aligned} \hat{s}_k &= \sum_{l=1}^L a_{kl}^* \check{s}_{kl} \\ &= \underbrace{\left( \sum_{l=1}^L a_{kl}^* \mathbf{v}_{kl}^H \mathbf{h}_{kl} \right)}_{\text{Effective channel}} s_k + \sum_{l=1}^L a_{kl}^* \left( \sum_{\substack{i=1 \\ i \neq k}}^K \mathbf{v}_{kl}^H \mathbf{h}_{il} s_i + \mathbf{v}_{kl}^H \mathbf{n}_l \right) \end{aligned} \quad (16)$$

The vector  $\mathbf{a}_k = [a_{k1} \dots a_{kL}]^T$  with weighting coefficients can be selected to maximize SE, but needs to be deterministic since the CPU does not know the channel estimates at Level 3. This approach is known as *Large-Scale Fading Decoding (LSFD)* in Cellular mMIMO [9]. From (16), we see that the signal of interest  $s_k$  is multiplied with the effective channel

$$\sum_{l=1}^L a_{kl}^* \mathbf{v}_{kl}^H \mathbf{h}_{kl} = \mathbf{a}_k^H \mathbf{g}_{kk} \quad (17)$$

with  $\mathbf{g}_{ki} = [\mathbf{v}_{k1}^H \mathbf{h}_{i1} \dots \mathbf{v}_{kL}^H \mathbf{h}_{iL}]^T$ . It is unknown at the CPU, but the average  $\mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}$  is deterministic and non-zero (if L-MMSE or MR is used) and thus can be assumed known. Based on this knowledge, the following SE is achievable.

**Proposition 2.** At Level 3, an achievable SE of UE  $k$  is

$$\text{SE}_k^{(3)} = \left( 1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left( 1 + \text{SINR}_k^{(3)} \right) \quad (18)$$

with the effective SINR given by

$$\text{SINR}_k^{(3)} = \frac{p_k |\mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}|^2}{\mathbf{a}_k^H \left( \sum_{i=1}^K p_i \mathbb{E}\{\mathbf{g}_{ki} \mathbf{g}_{ki}^H\} + \sigma^2 \mathbf{D}_k \right) \mathbf{a}_k - p_k |\mathbf{a}_k^H \mathbb{E}\{\mathbf{g}_{kk}\}|^2} \quad (19)$$

where  $\mathbf{D}_k = \text{diag}(\mathbb{E}\{\|\mathbf{v}_{k1}\|^2\}, \dots, \mathbb{E}\{\|\mathbf{v}_{kL}\|^2\})$ .

*Proof:* The result is proved by rewriting the system model (16) to fit [1, Cor. 1.3] and is omitted for space limitations. ■

The above SE holds for any combiner  $\mathbf{v}_{kl}$ , not only for L-MMSE combining in (15) or for MR with  $\mathbf{v}_{kl} = \hat{\mathbf{h}}_{kl}$  that was used in [7]. The structure of (19) makes it possible to find the deterministic weighting vector  $\mathbf{a}_k$  that maximizes the SINR.

**Corollary 2.** The effective SINR in (19) for UE  $k$  is maximized by

$$\mathbf{a}_k = \left( \sum_{i=1}^K p_i \mathbb{E}\{\mathbf{g}_{ki} \mathbf{g}_{ki}^H\} + \sigma^2 \mathbf{D}_k \right)^{-1} \mathbb{E}\{\mathbf{g}_{kk}\} \quad (20)$$

which leads to the maximum value

$$\begin{aligned} \text{SINR}_k^{(3)} &= p_k \mathbb{E}\{\mathbf{g}_{kk}^H\} \\ &\times \left( \sum_{i=1}^K p_i \mathbb{E}\{\mathbf{g}_{ki} \mathbf{g}_{ki}^H\} + \sigma^2 \mathbf{D}_k - p_k \mathbb{E}\{\mathbf{g}_{kk}\} \mathbb{E}\{\mathbf{g}_{kk}^H\} \right)^{-1} \mathbb{E}\{\mathbf{g}_{kk}\}. \end{aligned} \quad (21)$$

*Proof:* It follows by noting that (19) is a generalized Rayleigh quotient with respect to  $\mathbf{a}_k$ . ■

We note that Level 3 is an extension of the LSFD framework in [7], [9], [10], which has only been used in Cell-Free mMIMO along with MR combining. The signaling required at Level 3 is quantified as follows. Each AP needs to send  $(\tau_c - \tau_p)K$  complex scalars (i.e.,  $\check{s}_{kl}$  for all  $k$ ) per coherence block. In addition, the evaluation of (20) requires knowledge of the  $L$ -length complex vector  $\mathbb{E}\{\mathbf{g}_{kk}\}$  and  $L \times L$  matrices  $\mathbb{E}\{\mathbf{g}_{ki} \mathbf{g}_{ki}^H\}$  and  $\mathbf{D}_k$  for all  $k, i$ . Hence,  $KL + (L^2 K^2 + KL)/2$  complex scalars are needed in total, which can be a large number. These values are summarized in Table I.

### C. Level 2: Local Processing & Simple Centralized Decoding

Although the optimized LSFD step in Level 3 maximizes the SE, in practice the large number of statistical parameters must be estimated at the APs and it might not be feasible to share them with the CPU if the statistics vary with time. In that case, the CPU can form its estimate of the signal from UE  $k$  by taking the average of the local estimates. This yields

$$\hat{s}_k = \frac{1}{L} \sum_{l=1}^L \check{s}_{kl} \quad (22)$$

where  $\check{s}_{kl}$  is given in (14) and can be obtained by any local combining vector. Since this is equivalent to setting  $\mathbf{a}_k = [1/L \dots 1/L]^T$  in Proposition 2, the following is obtained.

**Corollary 3.** At Level 2, an achievable SE of UE  $k$  is

$$SE_k^{(2)} = \left(1 - \frac{\tau_p}{\tau_c}\right) \log_2 \left(1 + \text{SINR}_k^{(2)}\right) \quad (23)$$

with the effective SINR given by

$$\text{SINR}_k^{(2)} = \frac{p_k \left| \sum_{l=1}^L \mathbb{E} \{ \mathbf{v}_{kl}^H \mathbf{h}_{kl} \} \right|^2}{\sum_{i=1}^K p_i \mathbb{E} \left\{ \left| \sum_{l=1}^L \mathbf{v}_{kl}^H \mathbf{h}_{il} \right|^2 \right\} + \sigma^2 \sum_{l=1}^L \mathbb{E} \{ \|\mathbf{v}_{kl}\|^2 \} - p_k \left| \sum_{l=1}^L \mathbb{E} \{ \mathbf{v}_{kl}^H \mathbf{h}_{kl} \} \right|^2}. \quad (24)$$

This SE can be utilized along with any local combiner. If MR is used with single-antenna APs (i.e.,  $N = 1$ ), then Corollary 3 reduces to the case considered in [2]. The number of complex scalars to be exchanged per coherence block is the same as in Level 3, but no statistical parameters are needed. This is summarized in Table I.

#### D. Level 1: Small Cells

The simplest implementation level is when the signal from UE  $k$  is decoded using only the received signal from one AP [6, Sec. 4.2]. This decoding can be done locally at the AP, using its local channel estimates. Therefore, there is no need to exchange anything with the CPU.<sup>2</sup> This assumption essentially turns Cell-Free mMIMO into a small-cell network, but the macro diversity achieved by the AP-selection approach could potentially make it competitive compared to conventional Cellular mMIMO with larger cells.

Cell-Free mMIMO and small cells were compared in [2], [3] with  $N = 1$  and an AP selection based on the largest large-scale fading coefficient  $\beta_{kl}$ . Unlike [2], [3], we make no restrictions on  $N$  and let the AP that gives the highest SE to a specific UE be the one decoding its signal.<sup>3</sup>

**Corollary 4.** At Level 1, an achievable SE of UE  $k$  is

$$SE_k^{(1)} = \left(1 - \frac{\tau_p}{\tau_c}\right) \max_l \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{kl}^{(1)}\right) \right\} \quad (25)$$

where the instantaneous effective SINR for AP  $l$  is

$$\text{SINR}_{kl}^{(1)} = \frac{p_k |\mathbf{v}_{kl}^H \hat{\mathbf{h}}_{kl}|^2}{\sum_{\substack{i=1 \\ i \neq k}}^K p_i |\mathbf{v}_{kl}^H \hat{\mathbf{h}}_{il}|^2 + \mathbf{v}_{kl}^H \left( \sum_{i=1}^K p_i \mathbf{C}_{il} + \sigma^2 \mathbf{I}_N \right) \mathbf{v}_{kl}}. \quad (26)$$

The maximum in (26) is achieved with the L-MMSE combining in (15) and is given by

$$\text{SINR}_{kl}^{(1)} = p_k \hat{\mathbf{h}}_{kl}^H \left( \sum_{\substack{i=1 \\ i \neq k}}^K p_i \hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \sum_{i=1}^K p_i \mathbf{C}_{il} + \sigma^2 \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{kl}. \quad (27)$$

<sup>2</sup>In all the four levels, the  $K$  data streams need to be transmitted to the core network after decoding, which requires a backhaul load proportional to the sum SE. This is not included in Table I but is different for each level.

<sup>3</sup>We also noticed that the SE expression in [2, Eq. (47)] neglects the impact of pilot contamination during data transmission, thus it is an approximation.

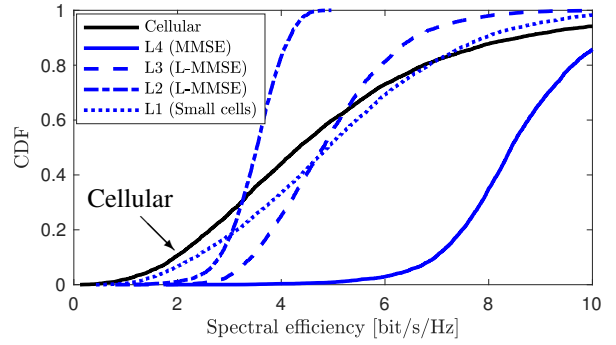


Fig. 2. Comparison of Cellular mMIMO with Cell-Free ( $L = 400$ ,  $N = 1$ ) when using MMSE or L-MMSE combining.

#### IV. CELL-FREE VERSUS CELLULAR mMIMO

In this section, we compare Cell-Free mMIMO, with the different cooperation levels, and Cellular mMIMO using the SE expression in [1, Sec. 4.1] and optimal multi-cell MMSE (M-MMSE) combining. The cellular network has 4 square cells in a  $1 \times 1$  km area, as in Fig. 1, with 100 antennas per BS. The cell-free network is deployed in the same area and has either 400 single-antenna APs (i.e.,  $N = 1$ ) or 100 four-antenna APs (i.e.,  $N = 4$ ). Hence, all the network configurations have the same number of antennas. To make a fair comparison, the APs are deployed on a square grid (instead of randomly as in [2], [3], [7]) and we use the same propagation model in all cases. It is based on the 3GPP Urban Microcell model in [11] with 2 GHz carrier frequency and

$$\beta_{kl} [\text{dB}] = -30.5 - 36.7 \log_{10} \left( \frac{d_{kl}}{1 \text{ m}} \right) + F_{kl} \quad (28)$$

where  $d_{kl}$  is the distance between UE  $k$  and AP  $l$  (computed as the minimum over different wrap-around cases) and  $F_{kl} \sim \mathcal{N}(0, 4^2)$  is the shadow fading. The shadowing terms from an AP to different UEs are correlated as [11, Table B.1.2.2.1-4]

$$\mathbb{E} \{ F_{kl} F_{ij} \} = \begin{cases} 4^2 e^{-\delta_{ki}/13 \text{ m}} & l = j \\ 0 & l \neq j \end{cases} \quad (29)$$

where  $\delta_{ki}$  is the distance between UE  $k$  and UE  $i$ . The second row in (29) accounts for the correlation of shadowing terms among two different APs, which is negligible since we always have at least 50 m between APs (notice that  $e^{-50/13} \approx 0.02$ ).<sup>4</sup>

The multi-antenna APs are equipped with half-wavelength-spaced uniform linear arrays and the spatial channel correlation is generated using the Gaussian local scattering model with  $15^\circ$  angular standard deviation [1, Sec. 2.6]. We use the same propagation models for the Cellular mMIMO case.

There are  $K = 40$  UEs, whereof ten are uniformly dropped in each cell and assigned to unique pilots with random indices. The same UE locations and pilot assignments are considered in the cell-free case, but the shadowing is generated independently. The transmit power is  $p_k = 100$  mW, bandwidth is 20 MHz,  $\sigma^2 = -96$  dBm,  $\tau_c = 200$ , and  $\tau_p = 10$ .

<sup>4</sup>The papers [2], [7] have established a standard model for evaluating Cell-Free mMIMO, but it is based on an older propagation model and has a shadow fading decorrelation distance that is  $10 \times$  larger than in the 3GPP model [11]. This is why we use the newer 3GPP model that we believe is more realistic.

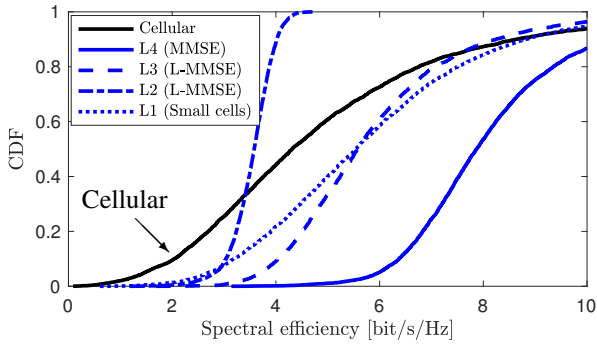


Fig. 3. Comparison of Cellular mMIMO with Cell-Free ( $L = 100$ ,  $N = 4$ ) when using MMSE or L-MMSE combining.

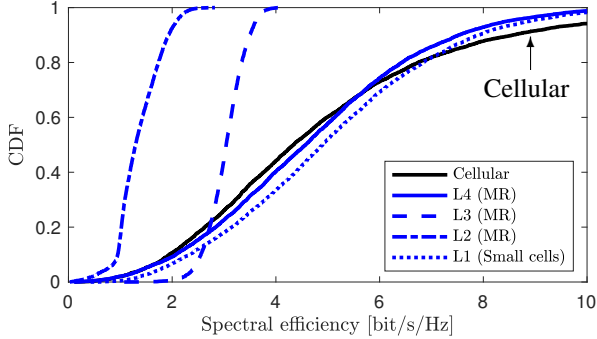


Fig. 4. Comparison of Cellular mMIMO with Cell-Free ( $L = 400$ ,  $N = 1$ ) when using MR combining.

Fig. 2 considers  $L = 400$  and  $N = 1$  and shows the cumulative distribution function (CDF) of the SE of a randomly located UE, when using MMSE or L-MMSE combining in the cell-free cases. At the 90% or 95% likely SE points, the cell-free cases perform according to their level: Level 4 is by far the best, while Level 1 gives the lowest SE but is anyway preferable as compared to Cellular mMIMO. Looking at the complete CDF curves, the situation is more complicated since the Level 1 and Cellular curves are crossing the Level 2 and Level 3 curves. Hence, UEs with better channel conditions get better performance with these cellular methods. However, Level 4 performs better than Cellular mMIMO for every UE.

Fig. 3 considers cell-free with  $L = 100$  and  $N = 4$  instead. The general trends are the same as in Fig. 2 but Level 4 loses in performance due to the reduced macro diversity, while Level 1 gains in performance since each AP can now suppress interference locally, by using its four antennas. In fact, Level 1 is now comparable to Level 2 for the weakest UEs and substantially better for the strongest UEs.

Finally, Fig. 4 considers the case  $N = 1$  and MR combining, which is the receiver processing advocated in the early papers on Cell-Free mMIMO. More precisely, Level 2 was considered in [2] and Level 3 in [7]. Compared to Fig. 2, the SE loss for Cell-Free mMIMO is very large due to the poor processing. In fact, Level 2 is outperformed by both small cells (Level 1) and Cellular mMIMO for every single UE. This is different from [2, Fig. 6], where Level 2 was the better choice for the majority of UEs. The reason is that we use a better UE association policy for the small cells, leading to more fair and competitive results. Not even Level 4 performs better than cellular.

#### A. A Quick Look at the Fronthaul Signaling Load

The reported results show that a Level 4 implementation is strongly preferred, otherwise, the SE gain over small cells is low. The counterargument might be that such an implementation would require much more fronthaul signaling than Level 2 and Level 3, but that is typically not the case. By using the numbers in Table I, Level 4 requires less signaling if

$$\frac{\tau_c N L}{(\tau_c - \tau_p) K L} = \frac{\tau_c}{\tau_c - \tau_p} \frac{N}{K} < 1. \quad (30)$$

Since  $\frac{\tau_c}{\tau_c - \tau_p} \approx 1$  and  $K \gg N$  are typical in Cell-Free mMIMO, Level 4 actually requires *much less* signaling. The reason is that the received data signals constitute a much larger number of scalars than the pilot signals. In fact, since  $K \geq N$  is typical, Level 2 and Level 3 actually increase the fronthaul signaling by processing the  $N$ -dimensional vector  $\mathbf{y}_l$  into the  $K$ -dimensional vector  $[\tilde{s}_{1l}, \dots, \tilde{s}_{Kl}]^T$ . Admittedly, this comparison assumes that all scalars are shared with infinite precision, while in practice it is plausible that the received pilot signals require higher bit-resolution than the data signals. On the other hand, [12] recently showed that the estimates can be compressed rather well in Cell-Free mMIMO.

## V. CONCLUSIONS

We compared Cell-Free and Cellular mMIMO and conclude that cell-free can provide vastly higher SE for all the UEs, but only if: (i) MMSE processing is used instead of MR; (ii) the received signals and estimates are sent to a CPU instead of being preprocessed at the APs (as previously advocated), since the latter greatly reduces the SE and increases the signaling.

## REFERENCES

- [1] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.
- [2] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017.
- [3] E. Nayeibi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Pricing and power optimization in cell-free Massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, 2017.
- [4] A. G. Burr, M. Bashar, and D. Maryopi, "Ultra-dense radio access networks for smart cities: Cloud-RAN, fog-RAN and "cell-free" massive MIMO," in *PIMRC, International Workshop of CorNer*, 2018.
- [5] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Asilomar Conference on Signals, Systems and Computers*, 2007, pp. 83–87.
- [6] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends® in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.
- [7] E. Nayeibi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *Asilomar SSC*, 2016, pp. 203–207.
- [8] H. Yang and T. L. Marzetta, "Energy efficiency of massive MIMO: Cell-free vs. cellular," in *VTC Spring*, 2018.
- [9] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, "Uplink interference reduction in large-scale antenna systems," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2194–2206, 2017.
- [10] T. Van Chien, C. Mollén, and E. Björnson, "Large-scale-fading decoding in cellular massive MIMO systems with spatially correlated channels," *IEEE Trans. Commun.*, 2019, to appear.
- [11] 3GPP, *Further advancements for E-UTRA physical layer aspects (Release 9)*. 3GPP TS 36.814, Mar. 2017.
- [12] D. Maryopi and A. G. Burr, "Few-bit CSI acquisition for centralized cell-free massive MIMO with spatial correlation," in *WCNC*, 2019.